

SYSTEMS AND METHODS FOR AUTOMATICALLY EQUALIZING AUDIO SIGNALS

The present invention generally relates to audio processing applications, and more particularly, to systems and methods for automatically equalizing audio signals based on information contained in the audio signal.

High quality audio systems typically utilize an equalizer to amplify or attenuate
5 selected frequency regions in order to enhance the perceived quality of the audio signal during playback. These equalizers may include a set of adjustable controls that allow the user to manually adjust the degree of amplification or attenuation of each frequency region until the desired result is obtained. Other approaches utilize a set of predetermined
10 equalizer settings that optimize playback for different types of music, such as pop, jazz or classical music. The user can then select the appropriate predetermined equalizer setting depending on the type of music being played.

Although these approaches can enhance the perceived quality of the audio signal, these approaches can produce sub-optimal results, especially for poor quality audio recordings with atypical spectral properties. For systems employing adjustable controls, for
15 example, the user typically must engage in a trial and error process in order to determine the optimal settings for each frequency region. Systems employing predetermined equalizer settings similarly require the user to switch between multiple equalizer settings in order to determine the setting that achieves the best results. These types of trial and error processes can lead to a frustrating listening experience for the user and may still produce sub-optimal
20 results depending on the technical capabilities of the user, the characteristics of the music involved and the quality of the audio recording.

Therefore, in light of the foregoing problems, there is a need for systems and methods for automatically equalizing audio signals. These systems and methods would preferably perform equalization based only on information contained in the audio signal,
25 thereby avoiding the need for the user to provide additional information about the audio recording.

Embodiments of the present invention alleviate many of the foregoing problems by providing systems and method for automatically equalizing audio signals. In one embodiment, spectral energy values for a plurality of frequency bands are measured by, for
30 example, measuring frequency components of at least a portion of the audio signal and equalizing the frequency components using previous or initial equalizer settings associated

with the frequency bands. Based on these measurements, spreaded energy values for each frequency band may be determined based on the measured spectral energy values and a spreading function, where the spreading function defines the spread of spectral energy across the frequency bands similar to the way it occurs within the auditory system. The ratio
5 between the unspreaded and spreaded energy in each frequency band is assumed to represent a measure of the audibility of the audio signal within that frequency band. Accordingly, this measure may be used to update equalizer settings for each frequency band by, for example, taking the previous or initial equalizer setting (in decibels) and subtracting a normalized ratio of the measured spectral energy value and the spreaded energy value for
10 that frequency band, where the normalized ratio may be determined by subtracting the above described ratio for that frequency band by an average ratio of unspreaded and spreaded energy values across all frequency bands under consideration to avoid an overall reduction in level. In this manner, the accumulation of adjustments to the equalizer settings results in slightly attenuating frequency bands that are clearly audible and slightly
15 amplifying frequency bands that are masked by other frequency bands. This process can enhance the perceptual quality of the audio signal by optimizing the information transmitted to the listener.

Other embodiments provide mechanisms for updating the equalizer setting so as to avoid switching artifacts and for further enhancing the perceptual quality of the audio
20 signal. For example, one embodiment multiplies the ratio of the spreaded and unspreaded energy values by a constant that is less than 1 in order to reduce the degree to which the equalizer settings can be adjusted in any iteration, thereby reducing switching artifacts. Other embodiments halve the equalizer settings used to amplify the audio signal in the decibel domain in order to provide a compromise between maximizing information transfer
25 and maintaining faithfulness to the original recording. Still further embodiments may configure the device to apply a graded equalization that allows for more adjustments in midrange frequency bands than in the far ends of the spectrum. By placing limits on the level of amplification at the far ends of the spectrum, these embodiments may reduce the level of noise or distortion perceived by the user, since very little signal is generally present
30 in these outer portions of the frequency spectrum.

An alternative or complementary embodiment for deriving a measure of audibility of the signal comprises directly comparing the ratio of the unspreaded energy and predetermined target value (which may be frequency dependent). Equalizer settings may

then be adjusted until the equalized energy distribution matches the target spectral energy function stored in memory. The embodiment can be used with many of the embodiments described above, such as embodiments directed to reducing switching artifacts, graded equalization for high, low and midrange frequency bands, and maintaining faithfulness to the original recording. This embodiment also has advantages in that it similarly does not require any additional information about the signal (other than the signal itself) and can adapt the equalizer settings to the spectral characteristics of the signal.

By automatically equalizing the audio signal based on the audio signal itself, embodiments of the present invention can significantly enhance perceived quality without the need for the user to engage in a difficult trial and error process. Furthermore, by deriving equalizer settings based on the spectro-temporal properties of the audio signal and adjusting the derived settings to maintain faithfulness to original recording, embodiments of the present can also achieve better equalization results than conventional approaches, especially for poor quality recordings with atypical spectral properties.

These and other features and advantages of the present invention will become more apparent to those skilled in the art from the following detailed description in conjunction with the appended drawings in which:

Figure 1 illustrates an exemplary block diagram of an audio playback system in accordance with one embodiment of the present invention;

Figure 2 illustrates an exemplary method in flowchart form for automatically equalizing audio signals in accordance with one embodiment of the present invention; and

Figure 3 illustrates an exemplary platform that may be used in accordance with embodiments of the present invention.

DETAILED EMBODIMENTS

Embodiments of the present invention provide systems and methods for automatically equalizing audio signals. The following description is presented to enable a person skilled in the art to make and use the invention. Descriptions of specific applications are provided only as examples. Various modifications, substitutions and variations of the preferred embodiment will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments and applications without departing from the scope of the invention. Thus, the present invention is not intended to be limited to the described and illustrated embodiments, and should be accorded the widest scope consistent with the principles and features disclosed herein.

Embodiments of the present invention may be used to enhance the perceived quality of audio signals during playback by automatically equalizing the audio signal. In one embodiment, an audio playback device, such as a CD player, MP3 player or digital power amplifier, may be configured with a single button (one key equalizer) for initiating
5 automatic equalization. When the button is pressed, the device automatically derives the proper equalizer settings that will enhance the perceived quality of the audio signal. Preferably, these equalizer settings are derived based on the spectro-temporal properties of the input signal only, thereby avoiding the need for the user to engage in a difficult and time-consuming trial and error process or provide additional information about the audio
10 signal.

In order to automatically derive the proper equalizer settings, the audio playback device may exploit psycho-acoustic principles. In this context, the audio signal may be analyzed to determine which parts of the audio spectrum that, at any moment in time, are clearly audible and which parts are masked by other spectral parts of the audio signal.
15 Depending on these measurements, the parts of the spectrum that are masked are slightly amplified, while the parts of the spectrum that are clearly audible are slightly attenuated. If the equalizer adjustments are accumulated over several tens of seconds, the equalizer settings will vary relatively slowly which helps avoid switching artifacts. Furthermore, if the measurements of masking versus non-masking are based on the equalized signals, then
20 the accumulation of adjustments will eventually reach an asymptote at some point, thereby producing steady state equalization settings.

The philosophy for using masking information to derive equalizer settings is based on maximizing the total information transmission to the listener. In other words, parts of the spectrum that are consistently (or very often) masked are increased in level until they
25 are not more often masked than other parts of the spectrum. However, it should be noted that simply selecting equalizer settings that maximize total information transmission can lead to undesirable results. From research on visual quality perception of color and contrast, for example, it was learned that optimum quality is reached when a compromise is made between maximum information transmission and faithfulness to a (visual)
30 reproduction.

For the particular problem of enhancing audio quality, there is generally not any additional information about the most faithful equalizer setting. Accordingly, for embodiments of the present invention, it is assumed that the audio signal itself is already set

such that it achieves the most faithful reproduction of the original. Therefore, in order to get the best possible audio quality that optimizes information transfer and achieves a faithful reproduction, a compromise is made between a flat equalizer setting and an adjusted equalizer setting based on masking information of the audio signal. A particularly good
5 compromise is reached when the equalizer adjustments resulting from maximizing the information transmission are halved in the decibel domain.

By providing a single button for automatically equalizing the audio signal, embodiments of the present invention can significantly enhance perceived quality without the need for the user to engage in a difficult trial and error process. Furthermore, by
10 deriving equalizer settings based on the spectro-temporal properties of the audio signal and adjusting the derived settings to maintain faithfulness to original recording, embodiments of the present can achieve better equalization results than conventional approaches, especially for poor quality recordings with atypical spectral properties. These and other advantages will readily apparent when viewed in conjunction with the following detailed embodiments.

15 Referring to Figure 1, an exemplary block diagram of an audio playback system in accordance with one embodiment of the present invention is illustrated generally at 100. As illustrated, the input signal to be equalized is applied to a segmentation windowing unit 110, which segments the input signal into a number of segments of, for example 10-50 ms, and applies an analysis window, such as a Hanning window, to each segment. The resulting
20 signal in each segment is then transformed to the frequency domain using an FFT and prefiltering unit 120. The FFT and prefiltering unit 120 also prefilters the frequency components with the reciprocal of an equal loudness curve to compensate for the non-equal perception of loudness at different frequencies. This process allows equalization to take into consideration the inherent perceptual characteristics of the human ear.

25 The output of the FFT and prefiltering unit 120 is then applied to an equalizing unit 140 that amplifies or attenuates each prefiltered frequency coefficient according to the current equalizer setting associated with each frequency band, $eq(i)$ (which for the first iteration are set flat). The equalized frequency components are then applied to an energy calculation unit 140 that determines the signal energy, $en(i)$, within each frequency band. It
30 should be noted that the width of each frequency band described above is preferably proportional to the equivalent rectangular bandwidth (ERB) scale. These ERB frequency bands generally provide a better estimate of auditory filter bandwidth than classical Bark scale models.

Once the frequency band energies are determined, a spreading unit 150 uses a spreading function to determine spreaded energy values for each frequency band, $sen(i)$. In this context, each spreaded energy value models the masking that occurs at a given frequency band number as a result of spectral energy at different frequency bands. For example, the spreaded energy values may be determined in accordance with the following formula:

$$sen(i) = \sum_j gfm(i,j)en(j) \quad (1)$$

where $sen(i)$ is the spreaded (masking) energy found in the i -th frequency band; $en(j)$ is the spectral energy in the j -th frequency band; and $gfm(i,j)$ is a spreading matrix that defines how the spectral energy distribution $en(j)$ is spread across auditory filters.

Based on the foregoing measurements of $en(i)$ and $sen(i)$, a signal-to-masking ratio unit 160 estimates the amount of masking at each frequency band by calculating the ratio between the spreaded and unspreaded energy values. An update unit 170 then uses the signal-to-masking ratio to update the equalization settings for each frequency band according to the following formula:

$$eq(i) = eq(i) - c*(smr(i) - \text{mean}(smr)) \quad (2)$$

where $eq(i)$ is the equalizer setting in decibels of the i -th frequency band; $smr(i)$ is the signal-to-masking ratio for the i -th frequency band and is equal to $en(i)/sen(i)$; $\text{mean}(smr)$ is the average signal-to-masking ratio across all frequency bands under consideration; and c is a constant for determining how fast the algorithm responds to changes in the signal-to-masking ratio. If c is set to a value of approximately 0.05, the algorithm gives a relatively slow adjustment of equalizer settings, but is sufficiently fast to adjust in a matter of tens of seconds. This relatively slow adjustment of equalizer settings helps avoid switching artifacts. Furthermore, by subtracting the signal-to-masking ratio for each frequency band by the average signal-to-masking ratio across all frequency bands under consideration, this process effectively normalizes the signal-to-masking ratio about the average so that some frequency bands would be amplified and some frequency bands would be attenuated, thereby avoiding an overall reduction in level.

Once the updated equalizer settings are determined, the update unit 170 then updates the equalizer settings within the equalizing unit 130. These updated equalizer settings are then used by the equalizing unit 130 for the next iteration of the algorithm. The update unit

170 also applies the updated equalizer settings to an equalizing postfilter 180 that amplifies or attenuates the frequency bands of the input signal according to the associated equalizer values. As noted above, using the updated equalizer settings provided by the update unit 170 may cause the output signal to depart from a faithful reproduction of the original signal, thereby causing annoying distortion. Accordingly, in order to provide a compromise between maximum information transfer and faithfulness of reproduction, embodiments of the present invention preferably adjust the equalizer settings provided by the update unit 170 by a compensation factor A. This compensation factor is preferably equal to 0.5 such that the equalizer settings used by the equalizing postfilter 180 to amplify or attenuate the input signal are halved in the decibel domain.

The equalizing that is obtained in this way will vary rather slowly and can be applied on a moment-to-moment basis without introducing artifacts. An alternative embodiment, however, may be configured to scan the entire musical recording or the entire CD using the processes described above and obtain an average value of the equalizer settings, eq(i). These average equalizer settings would then be used by the equalizing postfilter 180 for the entire recording or CD under consideration.

In order to further enhance the perceptual quality of the audio signal, other embodiments may replace the button for initiating automatic equalization with a knob that is configured to adjust the equalization strength. For example, the knob may be configured to increase or decrease the responsiveness of equalization (e.g., by adjusting the value of c in equation 2 above) and/or increase or decrease the level of information transfer relative to faithfulness to the original recording (e.g., by adjusting the value of the compensation factor A used by the equalizing postfilter 180).

Still other embodiments may configure the update unit 170 (or equalizing postfilter 180) to apply a graded equalization that allows for more adjustments in equalizer settings within midrange frequency bands than in the far ends of the spectrum. One exemplary grading function that works well is provided below:

$$A(n) = 0.25 - 0.25 * \cos(2 * \pi * n/N) \quad (3)$$

where n is the frequency band number; A(n) is the compensation factor, A, for frequency band n that is used by the update unit 170 (or equalizing and postfilter unit 180) to determine the equalizer settings used to amplify or attenuate the frequency bands of the audio signal; and N is the total number of frequency bands. Notably, for n=N and n=1, A(n)

is (almost) zero. Accordingly, most of the equalizing effect occurs at the midranges and less equalizing is present towards the lower or higher frequencies. By placing limits on the level of amplification at the far ends of the spectrum, these embodiments may reduce the level of noise or distortion perceived by the user, since very little signal is generally present in these outer portions of the frequency spectrum.

In an alternative or complementary embodiment, the update unit 170 may be configured to derive updated equalizer settings by directly comparing a ratio of the unspreaded energy, $en(i)$, and a predetermined target spectral energy distribution (which may be frequency dependent). Equalizer settings may be adjusted by the update unit 170 until the equalized energy distribution matches the target spectral energy function stored in memory. This embodiment can be used with many of the embodiments described above, such as embodiments directed to reducing switching artifacts, graded equalization at high, low and midrange frequency bands, and maintaining faithfulness to the original recording. This embodiment also has advantages in that it similarly does not require any additional information about the signal (other than the signal itself) and can adapt the equalizer settings to the spectral characteristics of the signal.

Referring to Figure 2, an exemplary method in flowchart form for automatically equalizing audio signals in accordance with one embodiment of the present invention is illustrated generally at 200. As illustrated, the exemplary method begins at step 210 by equalizing frequency components using the current equalizer settings associated with each frequency band. The equalized frequency components are then used to determine the energy within each frequency band at step 220. Based on these determined energy values, the spreaded energy values may be determined at step 230 using, for example, a spreading function as described in equation 1 above. A signal-to-masking ratio for each frequency band may then be determined at step 240 by dividing the measured energy value associated with the frequency band by the spreaded energy value associated with the frequency band. The equalizer setting for each frequency band may then be updated at step 250 by taking the previous equalizer setting and subtracting a normalized signal-to-masking ratio (e.g., the signal-to-masking ratio for the frequency band minus the average signal-to-masking ratio across all frequency bands under consideration) multiplied by a responsiveness factor, c , as described in equation 2 above. These updated equalizer settings may then be used to update the equalizer settings used at step 210 for the next iteration of the exemplary method. The updated equalizer settings may also be used at step 260 to set the equalizer settings used to

amplify or attenuate corresponding frequency bands of the input signal. In this regard, step 260 may multiply the updated equalizer settings determined in step 250 by a compensation factor A to maintain a certain degree of faithfulness to the original recording.

Referring to Figure 3, an exemplary platform that may be used in accordance with
5 embodiments of the present invention is illustrated generally at 300. As illustrated, the exemplary platform includes a microprocessor 310 operably coupled to a memory system 320 via a system bus 330. The memory system 320 may comprise a random access memory, a hard drive, floppy drive, a compact disk, or other computer readable medium, that stores computer instructions for an equalization module 340. The exemplary system
10 also includes an I/O interface 350 that is coupled to the microprocessor 310 and the memory system 320 in order to enable the system to input audio signals and output equalizer settings.

In operation, if the user presses an equalization button for initiating automatic equalization of the audio signal, the processor 310 uses the equalization module 340 to
15 process audio signals received from the I/O interface 350 and to determine the proper equalizer settings. These equalizer settings may then be output to the I/O interface 350 for use in amplifying the audio signal. If the equalization module 340 is implemented in
accordance with the principles described above, the exemplary system may be configured to automatically enhance the perceived quality of the audio signal by automatically deriving
20 the appropriate equalizer settings.

While the present invention has been described with reference to exemplary
embodiments, it will be readily apparent to those skilled in the art that the invention is not limited to the disclosed and illustrated embodiments but, on the contrary, is intended to cover numerous other modifications, substitutions and variations and broad equivalent
25 arrangements that are included within the scope of the following claims.